DOCUMENT RESUME

ED 435 191 FL 026 044

AUTHOR Kuruoglu, Guliz
TITLE Turkish Proficiency Tests: A National Model.
ISSN ISSN-1095-8096
PUB DATE 1999-00-00
NOTE 7p.
AVAILABLE FROM AATT Bulletin, Portland State University, Dept. of Foreign Languages and Literatures, P.O. Box 751, Portland, OR 97207.
PUB TYPE Journal Articles (080) -- Reports - Descriptive (141)
JOURNAL CIT AATT Bulletin; n23-24 p9-13 Spr-Fall 1999
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; Diagnostic Tests; Higher Education; *Language Proficiency; *Language Tests; Professional Associations; Program Descriptions; Rating Scales; Scoring; *Test Construction; *Turkish; Uncommonly Taught Languages
IDENTIFIERS Placement Tests

ABSTRACT

The process and results of a project to develop a Turkish language proficiency test are described. The project was undertaken by the American Association of Teachers of Turkic Languages, and involved development of two sample tests for intermediate and advanced levels in four skill areas (speaking, listening, reading, writing) using previously-developed proficiency guidelines. Both the format and the content of the tests are discussed. The instruments were designed as both diagnostic and placement tools for placement in college-level Turkish language courses, and incorporate authentic materials. In the listening and reading sections, examinees respond in English. Speaking and writing sections are answered in Turkish. Speech is evaluated on a five-point scale based on skills in grammar, comprehensibility, organization, vocabulary, and communication. Writing samples are evaluated for specific aspects of mechanics and content. Reading and listening skills are assessed on a less complex comprehension model. (MSE)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 435 191

TURKISH PROFICIENCY TESTS: A NATIONAL MODEL

GÜLIZ KURUOĞLU
UCLA

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Pelin Basci

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

FL 026044

TURKISH PROFICIENCY TESTS: A NATIONAL MODEL

Güliz Kuruoğlu
UCLA

Testing is an important element of language teaching. For years language teachers in the U. S. and elsewhere debated what the best way of testing a student should be as a part of classroom teaching. Applied linguists were interested in developing tests for use as research instruments. Similarly, for many years, teachers of Turkish had been discussing designing a suitable language test for diagnostic purposes for Turkish. After the work on Proficiency guidelines for Turkish have been finalized, teachers thought that it would be beneficial for the field to develop expertise on testing and assessment. A testing committee was formed and this committee started working on preparing tests which would be nationally viable. This paper discusses the work this committee has done in test design and development for Turkish. The paper focuses on four major points:

1. Objectives and expectations in the preparation of the aforementioned tests.
2. The purpose of these tests and where they could be used.
3. Organization of the tests.
4. Theoretical and practical aspects of scoring or assessment.

1. Objectives and expectations

Several years ago, AATT formed a Testing and Assessment committee in order to acquire some expertise on testing, to develop sample tests to assess Turkish language proficiency in all four skills (speaking, reading, writing and listening) for diagnostic and placement purposes, and to develop testing guidelines for all levels (from novice to superior). The committee members are: Pelin Başcı (PSU), Ender Creel (ILR), Mükrime Postacıoğlu (FSI) and myself. The committee first communicated by phone, setting some ground rules for the preparation and design of the tests. Later the Institute of Turkish Studies and UCLA came up with some funding which enabled the committee to meet in December 1997 at UCLA for two-and-a-half days to work on initial sample tests. During this meeting, the committee members discussed several issues related to testing. They focused mainly on developing two sample tests for intermediate and advanced levels. These tests were later sent to ARIT to be given to students who were going to go to Turkey to participate in the Summer Intensive Turkish program at Boğaziçi University. The committee did not have time and funding to prepare another test for the superior level. They felt at the time that it was more beneficial for teachers of Turkish to have intermediate and advanced level tests completed. The committee also did not have time to develop clear-cut guidelines for testing and assumed that, for the time being, the composition of these sample tests would substitute for the guidelines.

When the committee convened to develop the aforementioned sample tests, they looked at the issue of testing in a larger educational context and decided that there was a need to develop proficiency-based sample tests. Until recently, model Turkish language tests, intended for university students nationwide and prepared by a group of instructors and experts working at various universities and institutions, did not exist.[1] The only Turkish language tests which were used nationally were the ones prepared by ARIT and these were used in selecting university students at an advanced level to go to the Summer Intensive language program at Boğaziçi University.

For years, language instructors have been somewhat dissatisfied with the content and design of these ARIT tests. Members of AATT had been debating the issue of testing even during the preparation of Turkish Proficiency Guidelines and Language Learning

Framework for Turkish. Often discussions centered around how tests for Turkish language classes should look, what they should measure, and what their primary function should be. Teachers always have felt that something needed to be done to construct language tests that are valid, systematic, which compute reliably, and which are relevant to the curricula taught at universities. Teachers also felt that such an endeavor would require developing expertise in the field. With this in mind, the testing committee, when convened, decided to test students in four areas: speaking, listening, reading and writing.[2] In preparation for these tests the committee members consulted Turkish Proficiency Guidelines so that test questions comply with the levels indicated in these guidelines.

Designing the speaking test needed some research. After consulting with experts and discussing it with the committee members, we decided to ask speaking-proficiency questions in Turkish in the first part of this section in order to elicit personal information about the test taker: who they are, where they live and what they do. These questions, asked in Turkish, then were recorded on tape. The students would listen to the question on tape and record their answers (in Turkish) on a separate tape. In the second part of the speaking test, test takers are given a certain situation in which they are asked to perform in Turkish. The instructions for this part are also recorded on tape. Test takers also would record their own answers in Turkish. For the listening comprehension section, the committee members recorded dialogues among themselves on topics that are relevant to the particular proficiency level the tests intended to measure. For example, one of the topics which we included in the intermediate level test was 'getting an appointment from a doctor.' Since it was impossible to find authentic dialogues on this particular topic at the time, one of the committee members pretended to be a secretary at a doctor's office and the other person pretended to be the patient calling to get an appointment from the doctor. These two members sat down and recorded a dialogue. This dialogue was not prepared beforehand and those who participated in this dialogue made it up as they talked. In this sense the dialogues were as close to being authentic as possible. Several such dialogues were taped during the meeting. These tapes were later transcribed and re-recorded at the UCLA language lab in order to improve the sound quality and to eliminate unnecessary background noise.

Reading has always been an important part of testing a language. In proficiency testing, test takers are usually given short authentic passages to read. In the intermediate level, the students were given simple biographies, simple news items, and simple ads taken from newspapers and magazines. They were then asked to summarize the contents of the test by listing at least five points in the text. In the advanced test, the reading passages consist of short news items and a short letter taken from newspapers and magazines. Then the test taker is asked to summarize, in English, the content of the passage by answering all relevant questions such as who, what, when, how, and why.

Writing is one of the most difficult language skills to master. Testing writing proficiency "has a lot in common with the testing of speaking proficiency. First, a ratable sample must be elicited, then that sample must be scored holistically."[3] In order to elicit ratable samples for Turkish, students in intermediate and advanced levels were asked to write short compositions. The aim of the writing test was to get a good ratable sample from the test taker.

2. The purpose

Tests, in order to be useful, must be developed with a specific purpose, a particular group of test takers, and a specific language use domain (i.e. situation or context in which test takers will be using the language outside the test itself) in mind.[4] The committee's main purpose in preparing the aforementioned test was to provide samples

for diagnostic and placement purposes for Turkish language classes at institutions of higher learning in the U.S. Our intention was to prepare tests which could be used as samples so that instructors could develop similar tests at their home institutions to measure their students' proficiency. A second reason was to provide ARIT a sample set of tests as models for selecting students going to the Boğaziçi Program in İstanbul every year. Our main goal is, of course, to measure university students' level of competency in Turkish and to determine at what proficiency level they may perform when they go to Turkey to do doctoral research or to further their proficiency in the language.

Evaluating the overall usefulness of a given test is essentially subjective, and these particular tests reflect value judgments of the testing committee who prepared the tests. We had decided that these tests should be more proficiency-oriented. Since proficiency-oriented tests often use authentic materials in order to assess student competency, we, as the testing committee, tried to use authentic materials as much as possible.

3. Organization of the Tests

The tests are divided into four sections: speaking, listening, reading and writing. Students are required to finish each section in thirty minutes, and the whole exam takes two hours to complete. Each section is worth 25 points (total 100 points), and starts with clear directions to the test taker followed by questions. In speaking and writing sections, test takers are asked to provide speaking and writing samples in the target language, that is, in Turkish. In the listening and reading sections, test takers are asked to respond to questions in English. In the listening section, the test taker listens to a dialogue in Turkish. Then s/he answers comprehension questions in English. Similarly, in the reading section, test takers give a summary of the passage they read in English. The reasoning behind answering questions in English in listening and reading sections is to prevent test takers from using clues from the target language to answer these questions without properly comprehending the passage. Therefore this test divides language use into two distinct areas: productive and receptive. Speaking and writing sections of the test attempt to measure the productive language skills. These sections require students or test takers to create with the target language, and are intended for testers to elicit language samples in Turkish from test takers for assessment. With these samples in the target language, teacher or tester may observe how students or test takers use the language to express, interpret, or negotiate intended meanings, or simply create with the language. Listening and reading tests mostly measure receptive skills. In these sections, test takers do not create with the language but they utilize their knowledge of vocabulary and grammatical structures to comprehend texts or units of language.[5]

4. Assessment

The results of language tests are most often reported as numbers or scores. These scores then are used in making decisions. Sometimes this decision involves passing or failing a student, placing the student in a class appropriate to his level, or it may help instructors to rank and select students for placement purposes. Methods used to arrive at these scores are a crucial part of the measurement process. In the preparation of Turkish tests, deciding on the principles of assessment was one of the major tasks the testing committee had undertaken. For many years, even the most experienced teachers of Turkish have been discussing the problem of how to grade writing samples and speaking samples appropriately in the classroom or for placement purposes. ACTFL uses a global type of scoring and assigns a level, such as novice, intermediate, advanced, or superior to test takers. This method is quite useful in assessing the language competency of a student. However, grading test takers as intermediate or advanced would not serve well for

teachers of Turkish in the language classroom where the teacher has to assign a certain grade to a language student. The aim of the committee in preparing tests was both to evaluate these tests qualitatively and quantitatively. In other words, they both wanted to find a way to determine the degree of proficiency of the student in using Turkish, and to assign a number score to these tests. In order to solve this problem, after many discussions, the committee decided on the following solution:

In the speaking section, the test taker's speech may be evaluated in five content areas. These are: *grammar*, *comprehensibility*, *organization*, *vocabulary*, and *communication*. Each of these content areas are assigned five points:

1 = no functional ability,
2 = limited ability in speaking,
3 = moderate ability,
4 =extensive ability,
5 =complete ability in speaking.

If a student did poorly, then s/he would receive (1) which implies that s/he has no functional ability. If s/he did extremely well, the teacher or tester would give him/her (5) which would indicate that the student has complete ability in communicating his/her message.

Writing samples may also be evaluated in a similar manner. The committee therefore recommended that the writing section be rated by scoring samples in two main categories: mechanics and content. These two basic categories would have the following sub-categories:

Mechanics	Content
-spelling	-organization (paragraphing)
-grammar usage	-relevance to topic and cultural awareness
-punctuation	-creativity/appeal to reader
	-range of syntax
	-richness of vocabulary/ expression

Each of the above points then are assigned numerical points. These grades are given below:

Mechanics =10 pts.
Content = 15 pts.
Total = 25 points.

mechanics
punctuation = 2pts.
grammar = 6 pts.
spelling = 2 pts.
Total = 10 pts.

content
organization = 3 pts.
relevance to topic = 3 pts.
creativity = 3pts.
range of syntax = 3 pts.
richness of vocabulary = 3pts.
Total = 15 pts.

The committee spent more time in devising a grading system for productive language skills, mainly for speaking and writing. The listening section of the tests is not assigned such an elaborate grading system but is evaluated on the basis of whether students or test takers answer the listening comprehension questions correctly or not. The grading of this section emphasized the test taker's ability to hear, understand, follow, and process instructions, speech, news, conversation, and recorded discussion from a variety of recorded sources suitable for the student's level. Since any ability to understand recorded sources was the main focus of the listening comprehension section of the tests, the committee decided that test takers would receive 20 points from listening to the items on the tape recorded for listening purposes and would receive 5 points according to his/her ability to understand questions in the speaking section.

The reading comprehension part of the tests requires test takers to read several reading passages and summarize them in

English. Test takers are then assigned a grade according to their ability to read, understand, follow and process the written material. The grading depends on whether the test taker can draw correct inferences from the reading material, get the main ideas, and interpret contextual clues correctly.

5. Conclusion

The tests which I discussed in this paper are the first efforts of their kind designed for programs in U. S. universities. In preparing these tests we not only planned to provide teachers samples they can utilize and copy, but we also wanted to start an ongoing debate about the kinds of language tests which would be appropriate to give to students. We do not claim what the testing committee has prepared is the "ideal" test for all purposes. Even experts admit that "....there is no such thing as the one 'best' test even for a specific situation."[6] The test we prepared might be appropriate in some instances but not so in some others. We are hoping to get input and new ideas from our colleagues to make them better. Other approaches and suggestions in our effort to develop our tests would be welcome.

NOTES

[1] Various government institutions and the Defense Language Institute all have their own guidelines and tests; however, such a nationwide effort to have a uniform Turkish language test designed for institutions of higher learning did not exist before this time.

[2] For more on testing also see J.P.B. Allen and Alan Davies, *Testing and Experimental Methods* (London: Oxford University Press, 1977) and Mary Finocchario and Sydney Sako, *Foreign Language Testing* (New York: Regents Publishing Company Inc., 1983).

[3] Jerry W. Larson and Randall L. Jones, "Proficiency Testing for the Other Language Modalities," in *Teaching for Proficiency, the Organizing Principle*, ed. Theodore V. Higgs (Lincolnwood: National Textbook Company, 1989), 134.

[4] Lyle F. Bachman and Adrian S. Palmer, *Language Testing in Practice* (Oxford: Oxford University Press, 1996), 18.

[5] Lyle F. Bachman and Adrian S. Palmer, *Language Testing in Practice* (Oxford: Oxford University Press, 1996), 68.

[6] Lyle F. Bachman and Adrian S. Palmer, *Language Testing in Practice* (Oxford: Oxford University Press, 1996), 6.